

# Large Language Models Are Reasoning Teachers

Year: 2022 | Citations: 420 | Authors: Namgyu Ho, Laura Schmid, Se-Young Yun

---

## Abstract

Recent works have shown that chain-of-thought (CoT) prompting can elicit language models to solve complex reasoning tasks, step-by-step. However, prompt-based CoT methods are dependent on very large models such as GPT-3 175B which are prohibitive to deploy at scale. In this paper, we use these large models as reasoning teachers to enable complex reasoning in smaller models and reduce model size requirements by several orders of magnitude. We propose Fine-tune-CoT, a method that generates reasoning samples from very large teacher models to fine-tune smaller models. We evaluate our method on a wide range of public models and complex tasks. We find that Fine-tune-CoT enables substantial reasoning capability in small models, far outperforming prompt-based baselines and even the teacher model in many tasks. Additionally, we extend our method by leveraging the teacher model's ability to generate multiple distinct rationales for each original sample. Enriching the fine-tuning data with such diverse reasoning results in a substantial performance boost across datasets, even for very small models. We conduct ablations and sample studies to understand the emergence of reasoning capabilities of student models. Our code implementation and data are available at <https://github.com/itsnamgyu/reasoning-teacher>.